



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/605,709	06/27/2000	Frederick J. Damerau	YOR9-2000-0324US1	3738

30743 7590 02/10/2005

WHITHAM, CURTIS & CHRISTOFFERSON, P.C.
11491 SUNSET HILLS ROAD
SUITE 340
RESTON, VA 20190

EXAMINER

AZAD, ABUL K

ART UNIT	PAPER NUMBER
----------	--------------

2654

DATE MAILED: 02/10/2005

Please find below and/or attached an Office communication concerning this application or proceeding.



UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
P.O. Box 1450
ALEXANDRIA, VA 22313-1450
www.uspto.gov

MAILED

FEB 10 2005

Technology Center 2600

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Application Number: 09/605,709
Filing Date: June 27, 2000
Appellant(s): DAMERAU ET AL.

Clyde R Christofferson
For Appellant

EXAMINER'S ANSWER

This is in response to the supplemental appeal brief filed on June 29, 2004.

Art Unit: 2654

(1) *Real Party in Interest*

A statement identifying the real party in interest is contained in the brief.

(2) *Related Appeals and Interferences*

The applicant stated that there are no other appeals and interferences which will directly affect or be directly affected by or have a bearing on the decision in the pending appeal is contained in the brief.

(3) *Status of Claims*

The statement of the status of the claims contained in the brief is correct.

(4) *Status of Amendments After Final*

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) *Summary of Invention*

The summary of invention contained in the brief is correct.

(6) *Issues*

The appellant's statement of the issues in the brief is correct.

(7) *Grouping of Claims*

Appellant's brief includes a statement that claims 1-3, 6 and 4, 5 do not stand or fall together and provides reasons as set forth in 37 CFR 1.192(c)(7) and (c)(8).

(8) *Claims Appealed*

The copy of the appealed claims contained in the Appendix to the brief is correct.

Art Unit: 2654

(9) Prior Art of Record

5,819,220

SARUKKAI et al.

10-1998

(10) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claims 1-6 are rejected under 35 U.S.C. 102 (b). This rejection is set forth in prior Office Action, mailed on March 29, 2004 and reproduce below for convenience.

As per claim 1, Sarukkai teaches, “an automated method for setting up an a natural language interface in a Web site (col. 3, line 56 to col. 4, line 7, particularly reads on “in the context of speech interfaces to the web, the invention dynamically makes use of information provided by links in a document or the current page of the source document being viewed”) comprising the steps of”:

“defining a hierarchy of topics into which individual documents or Web pages can be classified” (col. 7, lines 17-60, reads on Table 1, here links are defining a hierarchy of topics);

“generating a keyword index for those documents” (col. 7, lines 17-60, reads on “the information shown in the table was extracted automatically by a simple parsing JAVA program shown in Appendix 1. The set of words constituting the link referent can constitute a web triggered word set, and it would make sense to base the speech recognition search towards this set of words since it is likely that the user will utter them); and

“for each topic in the hierarchy, a set of n-grams to a topic in the topic hierarchy, which set of n-grams is distinctive to the topic and wherein the n-grams maybe sparse or non-sparse n-grams” (col. 9, lines 17-24; and col. 10, lines 16-24; particularly reads on “the concept of extracting web-triggered word set information depending on the context of the web pages recently viewed can also be implemented in other methods. One method would be to appropriately smooth/re-estimate n-gram language model scores using HTML sources of the documents recently viewed”).

As per claim 2 Sarukkai teaches, “wherein the step of generating a keyword index comprises the step of extracting sparse n-grams of keywords for each group of pages in the topic hierarchy” (col. 9, lines 19-22, and col. 10, lines 16-24; reads on “n-gram language model score using the HTML sources of the documents recently viewed”).

As per claim 3, Sarukkai teaches, “further comprising the step of optionally reviewing and editing the keyword index” (col. 6, lines 36-39, reads on “modify the appropriate language Model and/or acoustic model parameters dynamically in step 34, using the selected word-set list (step 32), to be used during the speech recognition search process”).

As per claim 4, Sarukkai teaches, “an automated method for setting up an instance of natural language interface in a web site (col. 3, line 56 to col. 4, line 7,

Art Unit: 2654

particularly reads on “in the context of speech interfaces to the web, the invention dynamically makes use of information provided by links in a document or the current page of the source document being viewed”) comprising the steps of:”

“automatically inducing a topic hierarchy by examining a structure of the Web site” (col. 7, lines 17-60, reads on Table 1, here links are defining a hierarchy of topics);

“creating rules from the n-grams, wherein each topic has associated rules that are used to decide if a new input document or query references the topic” (col. 7, lines 17-60, reads on “the information shown in the table was extracted automatically by a simple parsing JAVA program shown in Appendix 1. The set of words constituting the link referent can constitute a web triggered word set, and it would make sense to base the speech recognition search towards this set of words since it is likely that the user will utter them” and col. 8, lines 54-67).

“creating rules from the n-grams, wherein each topic has associated rules that are used to decide if a new input document or query references the topic; creating n-grams from pages in the Website that are associated with a topic in the topic hierarchy wherein the n-grams may be sparse n-grams or non-sparse n-grams” (col. 9, lines 17-23 and col. 10, lines 10-24).

As per claim 5, Sarukkai teaches, “wherein the step of creating rules for a classification engine is performed automatically and further comprising the optional step of manually editing the rules” (col. 10, lines 10-15, particularly reads on “building grammars dynamically involves a lot of computation overhead. The web-trigger

Art Unit: 2654

approach does not dynamically vary the vocabularies. The web triggered word set boosting just selectively alters the scores that are assigned to the different words, treating the web triggered word sets differently”).

As per claim 6, it is interpreted and thus rejected for the same reasons set forth in the rejection of claim 2.

(11) Response to Argument

The applicant stated at page 9 of the supplemental brief, “note that the term “sparse n-gram”, as defined and used in the disclosed and claimed invention, are sequences of tokens or words from the text where the tokens or words may or may not have other words between them. Perhaps the term “sparse n-gram” has confused the Examiner into thinking that the n-grams as used in the art of speech/voice recognition is relevant to the claimed invention. However, both the specification as filed and the foregoing explanation have made clear that the claimed invention is using the concept of n-grams in a different way than used in the art of speech/voice recognition. All that is meant is the more generic notion of a set (or sequence) of not necessarily adjacent tokens or words in the text. So for instance, in a document about mortgage loan applications, which has the phrase “mortgage loan application” as distinctive, one would presumably identify the phrase “mortgage loan” or even the noncontiguous phrase “mortgage application” as characteristic of the document”.

The examiner was not thinking about n-grams as used in the art of speech/ voice recognition, instead in the rejection the examiner was thinking of and applying n-grams to a language model. An n-gram language model can be used both for the spoken and written text language.

The applicant argues at page 10 of the supplemental brief, "Sarukkai simply does not deal with any of the topics addressed in the disclosed and claimed invention. The present invention and Sarukkai have in common use of the term "n-gram" but at a technical level these uses are quite distinct. For Sarukkai, "n-gram" means a sequence of tokens that are assigned probabilities within the context of a speech recognition system language model, which is irrelevant to the claimed invention. Many systems use common technologies, but even here the details of usage are very different. One cannot reasonably maintain that Sarukkai anticipates or teaches any features the claimed invention".

The examiner disagrees with above assertion because Sarukkai does deal with all the topics addressed in the disclosed and claimed invention (see claims rejection). The applicant does not disclose nor claim whether n-gram are within the context of a speech recognition language model or in the context of text recognition language model. Since they can be used by both contexts, the n-grams used in Sarukkai are not irrelevant to the claimed invention.

The applicant further argues, "Sarukkai does not mention using a taxonomy of topics let alone inducing a taxonomy. As the current invention is not about the specific use of the taxonomy or classification rules (this is covered in patent No. 6,567,805 cross- referenced as patent application Serial No. 09/570,788) and none of the cited references or patents mention this, it can be seen that they do not say anything relevant about this key part of the invention".

The examiner disagrees with the applicant's above assertion because Sarukkai teaches using taxonomy of topics at table 1 link address and at column 6, lines 1-17. One example is given by Sarukkai at Table 1, about CS department Home page at University of Rochester, under the CS department Home Page and other related topic information hierarchy is categorized. Similar examples hold for other university departments.

The applicant argues at Page 11 of the supplemental appeal brief, "nor does Sarukkai mention using so-called sparse n-grams in the manner used in the current invention, namely, in conjunction with documents and groups of documents associated with nodes or topics in an (induced) hierarchy to identify collocations or phrases that are characteristic of the associated document or group of documents. Nor does Sarukkai mention converting sparse n-grams or collocations into classification rules, whose use is described in the context of a classification-based natural language interface for the Web in patent No. 6,567,805 (cross-referenced as application Serial No. 09/570,788). It follows from this that Sarukkai does not deal in any way with the combination of these

Art Unit: 2654

methods nor is such combination implicit in Sarukkai. It certainly cannot be reasonably maintained, when this is understood, that the claimed invention is anticipated by Sarukkai".

The examiner notes that the applicant claimed language as follows "generating a keyword index for those documents; and for each topic in the hierarchy, associating a set of n-grams to a topic in the topic hierarchy, which set of n-grams is distinctive to that topic and wherein the n-grams maybe sparse or non-sparse n-grams". Here applicant claimed n-grams maybe sparse or non-sparse n-grams. Therefore, any kind of n-grams can be read the claimed limitations. However, Sarukkai teaches a "sparse n-gram" which reads on "a set of word selectively extracted from the Web page source that is being currently displayed by the browser" (col. 7, lines 65-66).

The applicant further argues at page 11 of the Supplemental Appeal brief, "Sarukkai does teach the use of n-gram language models. However, the teachings of Sarukkai are not applicable to the claimed invention because they are not directed toward the set-up of a natural language interface. Sarukkai instead teaches methods for dynamically altering language models according to word sets in the documents searched. In other words, the language model is adjusted in response to documents found in a search. The n-grams used by Sarukkai are used for speech recognition, as known in the art. For example, Sarukkai teach smoothing or re-estimating "n-gram language model scores..." (col. 9, lines 20-21, emphasis added), thereby implying that the n-grams are used for speech recognition. N-grams are extremely well known in the

Art Unit: 2654

art of speech recognition. By comparison, the n-grams employed in the present invention are created from documents to be searched, and the n-grams are stored as an index for searching. Hence, the n-grams in the present invention are used for very different purposes compared to the n-grams of Sarukkai".

The examiner notes that the applicant here stated that Sarukkai does teach the uses of n-gram language model. According to Webster's II New Riverside University Dictionary, "natural language" meaning "a human written or spoken language". The claimed limitation "set-up of a natural language interface" reads on Sarukkai's "the set of words constituting the link referent can constitute a web triggered word set, and it would make sense to bias the speech recognition search towards this set of words since it is likely that the user will utter them. This web triggered word set can be supplemented with additional command words, function words, and even other triggered words that are commonly used in conjunction with them" (col. 7, lines 17-26). It is clear from Sarukkai's above statement that web triggered word set is extracted from web page source that is being currently displayed from the browser for set-up of a natural language interface. So the examiner disagrees with the applicant's statement that Sarukkai use n-gram only for speech recognition, infact applicant himself at the beginning of the above quoted paragraph said that Sarukkai does teach the uses of n-gram language model. Therefore, the applicant is contradicting himself about Sarukkai's teaching. Both Sarukkai's and the applicant's invention are created for documents to be searched, and the n-grams are stored as an index for searching (col. 8, lines 3-11). Therefore both Sarukkai and the instant application's invention are for same purpose.

Art Unit: 2654

The applicant argues at pages 14 and 15 of supplemental appeal brief that Sarukkai does not define a hierarchy of topics.

The examiner explained above that Sarukkai teaches defining a hierarchy of topics at Table 1 link address and page 6, lines 1-17.

The applicant argues at page 15 of supplemental appeal brief that Sarukkai does not teach generating a key word index for those document.

The examiner disagrees with applicant's assertion because Sarukkai teaches the limitation at column 7, lines 17-27, column 8, lines 3-11 and column 9, lines 17-24, and as also explained above in the response to the arguments.

The applicant argues again at page 16 of the supplemental brief that Sarukkai does not teach for each topic in the hierarchy . . ."

The examiner disagrees the applicant's assertion. Sarukkai teach at Table 1 each topic in the hierarchy.

The applicant asserts at page 16 of supplemental appeal brief as per claim 2, "Claim 2: "wherein the step of generating" The Examiner asserts that this claim element reads on Sarukkai, at col. 9, lines 19-22, and col. 10, lines 16-24 ("n-gram language model score using the HTML sources of the documents recently viewed"). This is incorrect. The examiners juxtaposition of the quote from the current invention and the citation from Sarukkai is simply absurd. What does generating an ordinary

Art Unit: 2654

search keyword index have to do with "n-gram language model score..."? The answer is- there is no connection whatsoever".

The examiner believes that Sarukkai's statement is correct. Sarukkai's invention better explained in his specification how n-gram language model score related with the keyword index.

The applicant asserts at pages 16 and 17 of supplemental appeal brief as per claim 3, "The Examiner asserts that the claim limitation added in claim 3 reads on Sarukkai at col. 6, lines 36-39 ("modify the appropriate language Model and/or acoustic model paramters dynamically in step 34, using the selected word-set list (step 32), to be used during the speech recognition search process"). This also evidences confusion. The review and possible modification mentioned in the current invention is manual and it is sensible to do this, i.e., a person could do this because the output are classification rules that people can understand. This is not the case with parameters of language and/or acoustic models, as in Sarukkai. It does not make sense to think one would or could manually review and modify a language Model and/or acoustic parameters. Such parameters are necessarily done with statistical estimation techniques.

There is no confusion in the Sarukkai. Sarukkai teaches editing the keyword index of the appropriate language model.

The applicant asserts at page 17 of supplemental appeal brief as per claim 4, "Claim element: "creating rules from the n-grams ..." The Examiner asserts that this claim element reads on Sarukkai, col. 7, lines 17-60, and col. 8, lines 54-67. This is incorrect for two reasons. First, n-grams as used in speech modeling are distinct from the n-grams discussed in the present invention, as discussed above. Second, Sarukkai does not create rules of any kind from n-grams. Rather he is using words extracted from documents to smooth parameters in a language or acoustic model".

The examiner disagrees with applicant's assertion because as discussed above in the response to applicant's argument that applicant's invention and Sarukkai both use an n-gram language model. Sarukkai does have a rule to extract web-triggered word set, as indicated by equation 3.

The applicant asserts at page 18 of supplemental appeal brief as per claim 5, "Claim 5, the Examiner asserts that the limitation added in claim 5 reads on the Sarukkai teaching "wherein the step of creating rules for a classification engine ..." found at col. 10, lines 10-15. This is incorrect. First, Sarukkai does not use rules, which he is criticizing in col. 10, line 9. Cf. col 3, lines 5-7. Sarukkai's invention is meant to be an alternative to writing grammar rules. Cf. col 3 lines 47-54, viz., dynamically updating statistical models. Second, the grammar rules referred to/criticized in Sarukkai are distinct from the topic classification rules discussed in the present invention. The grammar rules for speech recognition are not topic classification rules at all. They are rules for recognizing grammatical phrases or patterns in language to constrain the

Art Unit: 2654

output of a speech recognizer so that it is grammatical and likely. This is completely unrelated to the concerns of the current invention.

The examiner notes that limitation of claim 5 is "5. The automated method for setting up a natural language interface in a Web site recited in claim 4, wherein the step of creating rules is performed automatically and further comprising the optional step of manually editing the rules". The limitation does not claim "wherein the step of creating rules for a classification engine ..."

The applicant asserts at page 18 of supplemental appeal brief as per claim 6, "Claim 6 the response to the Examiner's assertion is covered by the above response to the Examiner's assertion regarding the limitation contained in claim 2".

The response to the above argument is given above in the response to the applicant's argument about claim 2.

The applicant further asserts at page 18 of supplemental appeal brief, "overall the examiner seems to be make identifications based on the use of the same word, out of context (n-gram, keyword, search, rule and to arbitrarily juxtapose parts of the two unrelated inventions based on these superficial word identifications. As demonstrated above, this does not make a *prima facie* case for anticipation".

The examiner disagrees with the applicant's assertion because Sarukkai not only teaches same words, Sarukkai also teaches same technical point of the invention.

Art Unit: 2654

Sarukkai teaches to create a spoken language interface by selectively extracted keyword from a web page source currently displayed by the browser.

In view of above response, the examiner has met his burden with regards to the first criterion in order to establish a case of *prima facie* for anticipation.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,



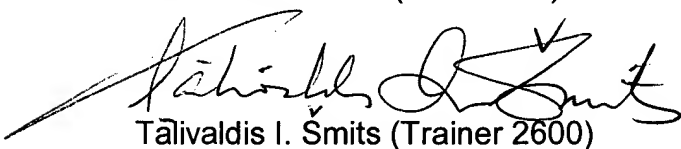
Abul K. Azad (Examiner)

February 4, 2005

Conferees



Richemond Dorvil (SPE 2654)



Tāivaldis I. Šmits (Trainer 2600)



Vijay Chawan (Primary Examiner 2654)

WHITHAM, CURTIS & CHRISTOFFERSON, P.C.
11491 SUNSET HILLS ROAD
SUITE 340
RESTON, VA 20190